Niveau(x) d'enseignement: 0 Numéro EDUQ : 8670389
Centre d'intérêt: 090

Bertrand, Richard G., 1949-.

L'invariance factorielle des méthodes de rotation orthomax : études théorique et expérimentale / Richard G. Bertrand, Normand Dufour. -- Sainte-Foy : INRS-Education, 1980.
28, [1] f.

Code: R-139. Bibliogr.: f. [29].

Vedette(s)-secondaire(s): Dufour, Normand * INRS-Education

Descripteur(s): Analyse de variables * Statistique

Identificateur(s) de nom: Kaiser, H. F.

Identificateur(s) de sujet: Analyse factorielle * Rotation orthomax

Après avoir expliqué ce qu'ils entendent exactement par "invariance factorielle", les auteurs montrent comment les travaux de H. F. Kaiser les ont menés à rechercher une condition suffisante à l'invariance factorielle pour les méthodes de rotation orthomax normalisées dans le cas de deux groupes de variables. Suite à ce résultat théorique, il relatent une étude expérimentale qui avait pour but d'investiguer dans quelle mesure ces mêmes méthodes de rotation pouvaient fournir une solution factorielle invariante dans le cas de deux groupes de variables fortement corrélées. Les résultats de l'étude montrent à quel point l'invariance factorielle peut être modifiée en ajoutant une ou deux variables de plus dans un groupe. Les résultats indiquent aussi l'importance d'un bon choix de variables lorsqu'on a l'intention d'entreprendre une analyse factorielle.

Type de document: r

R-139

L'INVARIANCE FACTORIELLE DES METHODES DE
ROTATION ORTHOGONALE ETUDES THEORIQUE ET
EXPERIMENTALE
PAR: RICHARD G. BEAUDRY ET NORMAND BEAUDIN
INRS-EDUCATION, QUEBEC, MAI 1980
DOCUMENT 80-130

EDUQ 8670389

L'INVARIANCE FACTORIELLE DES METHODES DE ROTATION ORTHOMAX: ETUDES THEORIQUE ET EXPERIMENTALE

Richard G. Bertrand
Normand Dufour

R-139

INRS-Education
2383, Chemin Ste-Foy
Ste-Foy, G1V 1T1

Mai 1980

L'INVARIANCE FACTORIELLE DES METHODES DE ROTATION ORTHOMAX: ETUDES THEORIQUE ET EXPERIMENTALE

INTRODUCTION

Depuis plus de vingt cinq ans, on a vu naître un très grand nombre de méthodes de rotation dites analytiques en analyse factorielle, parmi lesquelles on retrouve des rotations orthogonales telles quartimax (Neuhaus & Wrigley (1954)), varimax (Kaiser (1958)), equamax (Saunders (1962)) et des rotations obliques comme biquartimin (Carroll (1957)), promax (Hendrickson & White (1964)), orthotran (Hofmann (1978)). Ces méthodes de rotation ont toutes été mises sur pied afin que la solution factorielle finale, obtenue par rotation, puisse satisfaire à ce que Thurstone (1947) a appelé le principe de "structure simple". Or ce fameux principe a été introduit dans le but de résoudre le problème de manque d'invariance des facteurs. Kaiser (1958, p. 195) est très explicite quand il note:

> "The principle of simple structure may probably be considered incidental to the more fundamental concept of factorial invariance".

Après avoir expliqué ce que nous entendons exactement par "invariance factorielle" nous verrons comment les travaux de Kaiser (1958) nous ont mené à rechercher une condition suffisante à l'invariance factorielle pour les méthodes de rotation orthomax normalisées dans le cas de deux groupes de variables (phase théorique). A la suite de ce résultat théorique nous relaterons l'étude expérimentale qui avait pour but d'investiguer

dans quelle mesure ces mêmes méthodes de rotation pouvaient donner une solution factorielle invariante dans le cas de deux groupes de variables fortement corrélées (phase expérimentale).

PHASE THEORIQUE

Invariance factorielle

Qu'entendons-nous généralement par l'expression "invariance factorielle"? La meilleure réponse que nous avons trouvée à cette question se trouve dans Mulaik (1972, p. 337) lorsqu'il dit:

> "If different sets of variables representing the same domain of behavior are factor-analyzed, will the same fundamental factors be found? Or, if differently selected experimental populations are studied factor analytically with the same set of variables, will the same fundamental factors be found...?"

Il est clair qu'une réponse à l'une ou l'autre de ces questions est cruciale pour la validité d'une étude utilisant l'analyse factorielle. Thurstone (1947) et Kaiser (1958) sont même d'avis que l'invariance factorielle devrait être le critère ultime pour l'acceptation d'une méthode de rotation. Horst (1965, p. 418) quoique moins catégorique abonde aussi dans le même sens. Mulaik (1972) mentionne que le manque d'invariance factorielle peut mener à une situation où deux chercheurs indépendants entreprennent une analyse factorielle avec les mêmes variables et en arrivent à des facteurs fort distincts. De nombreux travaux ont déjà été entrepris en vue d'identifier clairement ce problème: Ahmavaara (1954), Meredith (1964), Jöreskog (1971), Bloxom (1972), ainsi que Thurstone (1947) et Kaiser (1958).

Dans cet article, précisons que notre intérêt se portera exclusivement sur le type d'invariance factorielle mentionné dans la première des deux questions de Mulaik citées plus haut. C'est donc dire que nous retiendrons comme source d'invariance factorielle, celle qui est déterminée par le changement de variables représentant le même domaine d'étude.

Du critère varimax normalisé au critère orthomax normalisé

Dès 1958, Henry Kaiser publiait un article définissant un critère analytique de rotation orthogonale, le varimax normalisé. Ce critère, très utilisé encore aujourd'hui, est considéré à juste titre comme un des meilleurs jamais produit (Rummel (1970) p. 392, Harman (1976) p. 297, Hofmann (1978)). Or Kaiser prouve mathématiquement que le varimax normalisé montre une invariance factorielle dans le cas particulier de deux groupes de variables colinéaires (i.e. m = 2 facteurs). Puis il donne un exemple pour un cas plus général où ce critère démontre, là aussi, une tendance à donner des facteurs invariants. Kaiser, toujours dans cet article, attribue l'invariance de son critère au processus de normalisation des coefficients des facteurs avant rotation et stipule que ni le quartimax ni le varimax non-normalisé ne possèdent cette propriété d'invariance. Ceci nous a incité à rechercher des conditions permettant l'invariance factorielle du quartimax normalisé et plus généralement des méthodes de rotation orthomax normalisées. Cette classe de méthodes de rotation est définie par Harman (1976, p. 299) de la façon suivante:

$$\Omega = \sum_{p=1}^{m} \left[\sum_{j=1}^{n} \left(\frac{a^2_{jp}}{h^2_j} \right)^2 - \frac{\gamma}{n} \left(\sum_{j=1}^{n} \frac{a^2_{jp}}{h^2_j} \right)^2 \right]$$

lorsque Ω atteint un maximum.

où : m : nombre de facteurs communs
n : nombre de variables
αjp : les coefficients des facteurs communs
h^2j : les communautés
γ : à déterminer

Pour $\gamma = 0$, on retrouve le critère quartimax (normalisé);
Pour $\gamma = 1$, on a le varimax (normalisé); $\gamma = \frac{m}{2}$ nous donne le critère équamax (normalisé), et $\gamma = \frac{1}{2}$ le biquartimax (normalisé).

Si on se place maintenant dans la situation particulière décrite par Kaiser (1958), où on a 2 groupes de variables colinéaires (i.e. m=2), on peut alors se poser la question suivante:
Quel est l'angle de rotation ϕ qui maximise Ω?

Si on définit: $\frac{a_{j1}}{h_j} = x_j$ et $\frac{a_{j2}}{h_j} = y_j$

et $X_j = x_j \cos\phi + y_j \sin\phi$

$Y_j = -x_j \sin\phi + y_j \cos\phi$

On trouve, en posant: $u_j = \left(\frac{a_{j1}}{h_j}\right)^2 - \left(\frac{a_{j2}}{h_j}\right)^2$

et $v_j = 2\left(\frac{a_{j1}}{h_j}\right)\left(\frac{a_{j2}}{h_j}\right)$

que (*) $\phi = \frac{1}{4} \arctan \left(\frac{2\left[n \Sigma u_j v_j - \gamma \Sigma u_j \Sigma v_j \right]}{n\Sigma\left(u_j^2 - v_j^2\right) - \gamma\left[\left(\Sigma u_j\right)^2 - \left(\Sigma v_j\right)^2\right]} \right)$

Si on fait $\gamma=1$ dans cette expression on retrouve l'équation (11) de Kaiser (1958 p. 194) i.e. le varimax, et si on fait $\gamma=0$ on retrouve $\tan 4\phi = \frac{D}{C}$, en reprenant les notations 13.24 de Harman (1976 p. 292), i.e. le quartimax.

Or, dans le cas où on a 2 groupes de variables colinéaires, Kaiser fait remarquer que toutes les valeurs v_j du premier (resp. du second) groupe sont égales et que toutes les valeurs w_j du premier (resp. du second) groupe sont égales.

Soit donc:

n_A : nombre de variables dans le premier groupe
n_B : nombre de variables dans le second groupe
v_a : valeur des v_j du premier groupe
v_b : valeur des v_j du second groupe
w_a : valeur des w_j du premier groupe
w_b : valeur des w_j du second groupe

Ainsi l'équation (*) devient

$$(**)\quad \phi = \frac{1}{4} \arctan \left(\frac{2\left\{(1-\gamma)\,(n_A^2 v_A w_A + n_B^2 v_B w_B) + n_A n_B \left[v_B w_B + v_A w_A - \gamma\,(v_A w_B + v_B w_A)\right]\right\}}{(1-\gamma)\left[n_A^2 (v_A^2 - w_A^2) + n_B^2\,(v_B^2 - w_B^2)\right] + n_A n_B \left[(v_A^2 - w_A^2) + (v_B^2 - w_B^2) - 2\gamma (v_A v_B - w_A w_B)\right]} \right)$$

Si $\gamma=1$ (varimax) dans (**), on retrouve exactement l'expression (12) de Kaiser (1958, p. 195) qui indique, comme on peut le voir et comme le souligne Kaiser, l'indépendance du critère normal varimax par rapport à n_A et n_B, et donc la propriété d'invariance.

Si $\gamma=0$ (quartimax), on obtient: $\phi = \frac{1}{4} \arctan \left(\frac{2\,(n_A v_A w_A + n_B v_B w_B)}{n_A (v_A^2 - w_A^2) + n_B\,(v_B^2 - w_B^2)} \right)$

En inspectant cette dernière expression, on se rend compte que si chacun des 2 groupes de variables a la même taille (i.e. $n_A = n_B$), l'angle de rotation ϕ devient indépendant de n_A et n_B puisqu'alors

$$\phi = \frac{1}{4} \arctan \left(\frac{2 (u_A v_A + u_B v_B)}{(u_A^2 - v_A^2) + (u_B^2 - v_B^2)} \right)$$

Cette condition (i.e. $n_A = n_B$) imposée au critère quartimax normalisé pour qu'il acquière la propriété d'invariance factorielle est d'autant plus intéressante qu'elle est généralisable pour γ quelconque.

En effet, si on fait $n_A = n_B$ dans (**), on obtient:

$$\phi = \frac{1}{4} \arctan \left\{ \frac{2n_A^2 \left[(1-\gamma)(u_A v_A + u_B v_B) + \left(u_B v_B + u_A v_A - \gamma(u_A v_B + u_B v_A) \right) \right]}{n_A^2 \left[(1-\gamma)\left((u_A^2 - v_A^2) + (u_B^2 - v_B^2) \right) + (u_A^2 - v_A^2) + (u_B^2 - v_B^2) - 2\gamma(u_A u_B - v_A v_B) \right]} \right\}$$

D'où le résultat général: Dans le cas où on a 2 groupes de variables colinéaires (i.e. $m = 2$), la condition $n_A = n_B$ est suffisante pour garantir l'invariance factorielle des méthodes de rotation orthomax normalisées.

PHASE EXPERIMENTALE

Description des groupes de variables étudiées

Ce résultat théorique démontré, nous avons voulu vérifier expérimentalement dans quelle mesure les méthodes de rotation orthomax normalisées montraient une invariance factorielle pour deux groupes de variables de même taille et fortement corrélées

(autant que possible 0,5 ou plus): ceci dans le but de s'éloigner le moins possible du cadre théorique. Trois ensembles de données ont été utilisés à cette fin. Le premier a été formé à partir du "Thurstone box problem" (Thurstone (1947)). Le second et le troisième ont été dérivés du projet de recherche sur les "Styles d'apprentissage" dirigé conjointement par le Dr. Horacio J.A. Rimoldi du CIIPME et le professeur Yves Bégin de l'INRS-Education.[1]

Les deux groupes de variables retenus dans le premier cas correspondent respectivement à un groupe de variables très corrélées avec la dimension "x" d'une part, et d'autre part à un groupe de variables très corrélées avec la dimension "y":

Groupe I	Groupe II
$V_{11} : x$	$V_{21} : y$
$V_{12} : xz$	$V_{22} : yz$
$V_{13} : x^2z$	$V_{23} : y^2z$
$V_{14} : xz^2$	$V_{24} : yz^2$
$V_{15} : x/z$	$V_{25} : y/z$
$V_{16} . 2x + 2z$	$V_{26} : \sqrt{y^2 + z^2}$

La matrice de corrélations de ces variables a été tirée de Thurstone (1947 p. 370) et se trouve au tableau I.

---

1 Nous tenons à remercier chaleureusement le professeur Yves Bégin et le docteur Horacio J.A. Rimoldi pour leur appui à cette recherche et pour leur permission concernant l'utilisation des données.

Dans le premier des deux ensembles de données du projet "Styles d'apprentissage" (cas où 184 sujets argentins et canadiens sont testés), les deux groupes de variables considérés sont:

Groupe I	Groupe II
V_{11} : "Sommation"	V_{21} : "Obj. con. 33"
V_{12} : "Synonimes"	V_{22} : "Obj. géo. 31"
V_{13} : "Divisions"	V_{23} : "Obj. géo. 33"
V_{14} : "Rel. Quant."	V_{24} : "Ve. Qual. 33"
V_{15} : "Mat. Figure"	V_{25} : "Ve. Qt. 33"
V_{16} : "Copies"	V_{26} : "Abstra 33"

Pour le dernier ensemble de données, lui aussi provenant du projet "Styles d'apprentissage" (cas où 88 sujets canadiens sont testés) on a choisi les variables suivantes:

Groupe I	Groupe II
V_{11} : "Sommation"	V_{21} : "Obj. con. 33"
V_{12} : "Synonimes"	V_{22} : "Obj. géo. 31"
V_{13} : "Gottschald"	V_{23} : "Obj. géo. 33"
V_{14} : "Divisions"	V_{24} : "Ve. Qual. 33"
V_{15} : "Refram"	V_{25} : "Ve. Qt. 33"
V_{16} : "Prob. ana 2"	V_{26} : "Abstra 33"

Les tableaux II et III nous montrent les matrices de corrélation correspondant respectivement au deuxième et au troisième ensembles de données.

TABLEAU I

Matrice de corrélations des deux groupes de variables issus du "Thurstone box problem"

	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}	V_{16}	V_{21}	V_{22}	V_{23}	V_{24}	V_{25}	V_{26}
V_{11}	1											
V_{12}	76	1										
V_{13}	89	95	1									
V_{14}	61	97	87	1								
V_{15}	59	-03	21	-24	1							
V_{16}	72	99	93	97	-08	1						
V_{21}	24	37	33	35	-03	35	1					
V_{22}	26	69	54	74	-42	69	79	1				
V_{23}	24	59	47	61	-31	58	89	96	1			
V_{24}	26	75	58	83	-50	75	67	97	90	1		
V_{25}	-06	-40	-30	-50	43	-44	56	02	23	-17	1	
V_{26}	29	68	54	72	-37	68	80	97	94	93	05	1

Note: Par souci de clarté on a omis le zéro et la virgule des coefficients de corrélation: par exemple 76 veut dire 0,76.

TABLEAU II

Matrice de corrélations des deux groupes de variables du projet "Styles d'apprentissage" (cas 184 sujets)

	V11	V12	V13	V14	V15	V16	V21	V22	V23	V24	V25	V26
V11	1											
V12	52	1										
V13	68	58	1									
V14	71	46	47	1								
V15	54	42	31	59	1							
V16	53	51	60	40	37	1						
V21	06	19	22	09	00	15	1					
V22	03	13	20	01	01	21	42	1				
V23	14	31	29	13	12	22	59	27	1			
V24	09	20	22	11	04	22	55	23	49	1		
V25	11	19	23	17	03	15	41	17	35	59	1	
V26	16	30	33	08	-01	22	42	19	35	45	54	1

TABLEAU III

Matrice de corrélations des deux groupes de variables du projet "Styles d'apprentissage" (cas 88 sujets)

	V11	V12	V13	V14	V15	V16	V21	V22	V23	V24	V25	V26
V11	1											
V12	60	1										
V13	51	62	1									
V14	77	63	47	1								
V15	50	64	46	51	1							
V16	58	57	45	65	61	1						
V21	04	15	05	15	08	02	1					
V22	-01	06	-09	06	05	11	39	1				
V23	14	28	12	20	14	09	67	22	1			
V24	-04	23	08	07	08	01	64	22	52	1		
V25	06	22	04	16	05	14	44	11	38	59	1	
V26	26	36	17	27	25	29	37	04	32	34	44	1

Procédure[1]

Deux méthodes d'extraction des facteurs ont été utilisées: l'analyse en composantes principales et l'analyse factorielle classique. Dans ce dernier cas, deux estimés de communauté ont été considérés soit MAXROW (i.e. la valeur absolue de la corrélation la plus élevée de chaque colonne de la matrice de corrélation) et SMC (i.e. le coefficient de corrélation multiple au carré). De plus on a employé 5 méthodes de rotation. Ces méthodes furent obtenues en faisant varier le terme γ dans l'expression Ω de la classe des rotations orthomax (établie plus haut) comme suit:

$\gamma = 0$ (quartimax), $\gamma = 0,25$, $\gamma = 0,5$ (biquartimax)
$\gamma = 0,75$, $\gamma = 1$ (varimax).

Pour chaque ensemble de données, on a procédé de la façon suivante. On a d'abord choisi, dans chacun des deux groupes, deux variables (parmi les plus corrélées). Puis, on a analysé la matrice de corrélation 4 x 4 correspondante en employant les 3 estimés de communautés (1, MAXROW, SMC), obtenant ainsi 3 couples de facteurs avant rotation. Enfin pour chaque couple de facteurs avant rotation on a généré 5 couples de facteurs (après rotation) en utilisant les 5 méthodes de rotation décrites plus haut.

---

1 La procédure informatique est la suivante: nous avons utilisé dans le "package" BMDP (Biomedical Computer Programs P-Series, University of California Press, 1977), le programme "Factor analysis" 4M, de telle sorte que l'on obtienne toujours 2 facteurs. En ce qui concerne l'extraction de facteurs pour le modèle de l'analyse factorielle classique, le programme procède par une procédure itérative sur les communautés placées dans la diagonale de la matrice de corrélation. Cette procédure est contrôlée par un critère de convergence sur les communautés obtenues.

Après cette analyse complète de la matrice 4 x 4, on a ajouté une variable (parmi les plus corrélées avec les deux premières) dans chaque groupe et on a analysé la matrice de corrélation 6 x 6 correspondante de la même façon que la matrice 4 x 4 décrite plus haut.

On a procédé ainsi en ajoutant successivement une variable à chaque groupe et en analysant les matrices 8 x 8, 10 x 10 et 12 x 12 correspondantes.

Le tableau IV donne un exemple de ce que nous avons obtenu en procédant de la sorte. La partie de gauche du tableau correspond aux coefficients du facteur 1 et la partie de droite aux coefficients du facteur 2.

TABLEAU IV

Evolution des coefficients des facteurs dans le cas du "Thurstone box problem"
(Méthode d'extraction: analyse en composantes principales, méthode de rotation: γ=0 (quartimax))

	Facteur 1					Facteur 2				
	4	6	8	10	12	4	6	8	10	12
V_{11}	08	06	03	02	01	97	94	91	89	86
V_{21}	96	93	90	88	87	12	09	07	03	01
V_{13}	26	27	27	28	27	94	95	95	94	94
V_{23}	96	96	97	97	96	20	22	22	19	20
V_{12}		40	42	44	43		87	89	89	90
V_{26}		91	92	93	93		31	32	31	32
V_{16}			41	44	43			88	88	89
V_{22}			93	94	94			31	30	31
V_{14}				51	51				81	83
V_{24}				88	88				39	41
V_{15}					-49					22
V_{25}					23					-54

$V_1 = 57,64 \times 10^{-6}$
$I_1 = 0,99994236$

$V_2 = 86,23 \times 10^{-6}$
$I_2 = 0,99991377$

Chaque ligne du tableau nous indique la variation du coefficient (obtenu après rotation) suivant qu'il provient d'une analyse de la matrice de corrélation 4 x 4, 6 x 6, 8 x 8, 10 x 10 ou 12 x 12 correspondante. L'ordre des lignes correspond à l'ordre d'entrée des variables (de chaque groupe) dans la matrice analysée. Ce tableau a été formé à partir d'analyses successives de sous-matrices de la matrice de corrélation initiale (tableau I). On a tout d'abord choisi deux variables pour chaque groupe, V_{11} et V_{13} pour le groupe I, V_{21} et V_{23} pour le groupe II (V_{13} étant la variable la plus corrélée avec V_{11} et V_{23} la plus corrélée avec V_{21}). Puis on a fait une analyse en composantes principales de la matrice de corrélation 4 x 4 correspondante $\begin{pmatrix} 1 & & & \\ 24 & 1 & & \\ 89 & 33 & 1 & \\ 24 & 89 & 47 & 1 \end{pmatrix}$, suivie d'une rotation orthomax normalisée ($\gamma = 0$). Ainsi le nombre 08 (resp 97) correspond au coefficient du facteur 1 (resp. facteur 2) pour la variable V_{11} après une analyse en composantes principales sur la matrice 4 x 4, suivie d'une rotation orthomax normalisée ($\gamma = 0$). La même interprétation vaut pour les variables V_{21}, V_{13}, V_{23}. Puis les couples de variables (V_{12}, V_{26}), (V_{16}, V_{22}), (V_{14}, V_{24}) et enfin (V_{15}, V_{25}) ont été introduits successivement en vue de montrer l'évolution des coefficients. Il est alors possible de comparer les coefficients de l'un à l'autre des deux facteurs en centrant notre attention sur les diverses colonnes de la partie gauche ou de la partie droite du tableau IV.

Indice d'invariance factorielle

Afin de rendre ce processus de comparaison des coefficients plus analytique nous avons construit un indice d'invariance

factorielle Ip qui a été calculé pour chaque facteur p. On définit

$$Ip = 1 - Vp$$

$$\text{où } Vp = \frac{\frac{1}{4} \sum_{n \in X} \sum_{j=1}^{n+\delta} (\alpha_{j(n+\delta)p} - \alpha_{j(n+\delta+2)p})^2}{d} \quad , p = 1, 2$$

tel que $X = \{2x + 4 / x \in \mathbb{N}, 2x + 6 + \delta \leq v\}$,

v : nombre total de variables considérées,

$\delta = n_1 - n_2$, n_1 (resp. n_2) étant le nombre de variables du premier (resp. second) groupe,

$\alpha_{j(n+\delta)p}$: coefficient du tableau de "l'Evolution des coefficients des facteurs" qui correspond à la $j^{ième}$ variable (ligne) et au $p^{ième}$ facteur, obtenu de l'analyse de la matrice n x n pour un "δ" spécifié au préalable.

d : nombre de différences du numérateur de Vp.

Le terme 1/4 est un facteur de pondération qui sera expliqué dans le paragraphe qui suit.

L'indice Ip varie dans un intervalle de 0 à 1; en effet sa valeur maximale est 1, dans le cas où toutes les différences entre coefficients sont nulles, i.e. Vp = 0.
La valeur minimale de Ip est 0, dans le cas où les différences entre coefficients sont maximales et donc égales à 2 en valeur

absolue: à ce moment $Vp = \frac{1}{4} \frac{d.\ 2^2}{d} = 1.$

Ainsi plus l'indice d'invariance factorielle Ip augmente plus on dénote une invariance factorielle. Ou encore plus l'indice Vp, de "variation factorielle" diminue plus on dénote une invariance factorielle.

En traçant les courbes de Vp[1] en fonction de γ , tel que par exemple au tableau V, on pourra se rendre compte de l'invariance factorielle relative d'une méthode de rotation par rapport à une autre.

DISCUSSION

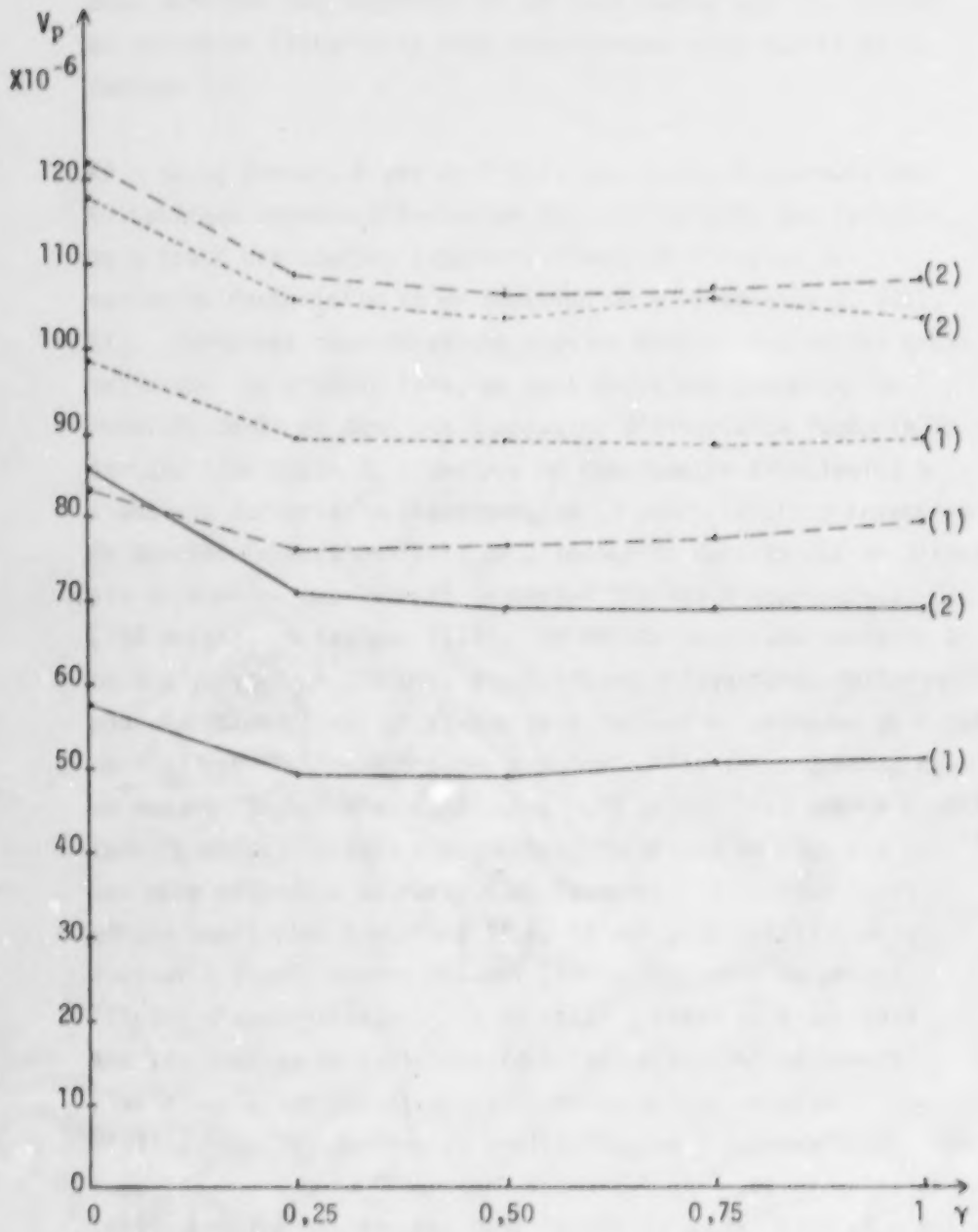
Pour chaque méthode d'extraction de facteurs et pour chaque méthode de rotation de chaque ensemble de données on obtient un tableau comme le tableau IV. On a donc 3 x 5 x 3 = 45 tableaux distincts qui révèlent l'évolution des coefficients des facteurs. Par exemple au tableau IV, on se rend compte que, dans le cas du "Thurstone box problem", en utilisant comme méthode d'extraction de facteurs, l'analyse en composantes principales, et comme méthode de rotation le quartimax (γ=0), les coefficients des facteurs changent très peu. En effet, pour le facteur 1, l'indice de variation factorielle V_1 est aussi petit que 57, 64 x 10^{-6} et pour le second facteur, V_2 = 86, 23 x 10^{-6} . Le même phénomène peut être observé dans le cas des deux autres ensembles de données si on utilise la même méthode d'extraction des facteurs et la même méthode

---

1 L'indice Vp a été préféré à l'indice Ip car il semble plus facile de voir la différence entre par exemple 57, 64 x 10^{-6} et 1923 x 10^{-9} qu'entre 0,99994236 et 0,99807700

TABLEAU V

Courbes (approximées) de l'évolution de V_p en fonction de γ
(Thurstone box problem)



Légende: ——— : PCA
– – – : PFA Maxrow
----- : PFA SMC
(1) : facteur 1
(2) : facteur 2

Remarque: Les indices V_p et I_p ont été calculés en tenant compte de trois chiffres significatifs des coefficients des facteurs.

de rotation. Les tableaux VI et VII en témoignent. Dans ces deux derniers cas toutefois on se rend compte que les indices de variation factorielle sont sensiblement plus élevés qu'au tableau IV.

Afin de se donner, à peu de frais, une vision d'ensemble des 45 tableaux donnant l'évolution des coefficients des facteurs, on a tracé des courbes (approximatives) de l'indice de variation factorielle Vp en fonction de γ (tableaux V, VIII, IX). Certaines considérations peuvent être tirées de ces trois tableaux. En premier lieu, on peut noter une augmentation sensible de Vp et donc une diminution d'invariance factorielle lorsque l'on passe de l'analyse en composantes principales à l'analyse factorielle classique, et ce pour les trois ensembles de données (exception faite de l'indice V2 dans le cas de l'exemple de données que fournit le projet "Styles d'apprentissage" (184 sujets) au tableau VIII). On dénote aussi une tendance à ce que Vp soit un peu plus élevé (moins d'invariance factorielle) pour quartimax (γ=0) et stable pour les autres méthodes de rotation, dans le cas du "Thurstone box problem". Alors que dans le cas du projet "Styles d'apprentissage" (88 sujets), Vp semble croître (sensiblement) lorsque γ augmente. Observons de plus que pour une même méthode d'extraction de facteurs, le facteur 1 est généralement plus invariant (i.e. V1 est plus petit) que le facteur 2 (sauf, encore une fois, dans l'exemple du projet "Styles d'apprentissage" (84 sujets)). Enfin on s'aperçoit que les indices de variation factorielle Vp sont nettement plus élevés, impliquant donc une moins grande invariance factorielle, pour les données du projet "Styles d'apprentissage" que pour les données du "Thurstone box problem". Les corrélations sensiblement moins élevées dans le cas du projet "Styles d'apprentissage" semblent devoir être la source de ce phénomène.

TABLEAU VI

Evolution des coefficients des facteurs dans le cas du projet "Styles" (184 sujets)
(Méthode d'extraction: analyse en composantes principales, méthode de rotation: γ=0 (quartimax))

	Facteur 1					Facteur 2				
	4	6	8	10	12	4	6	8	10	12
V11	93	93	90	90	88	05	02	08	05	03
V21	02	04	00	01	02	90	86	83	84	80
V14	92	85	86	83	80	07	03	05	01	02
V23	10	13	12	13	16	89	82	77	75	70
V13		79	70	73	74		25	34	33	30
V24		07	04	06	06		81	80	76	79
V15			76	73	72			-06	-08	-09
V26			09	12	13			70	66	70
V16				68	69				26	22
V22				03	05				54	48
V12					71					26
V25					08					72

$V_1 = 247{,}27 \times 10^{-6}$
$I_1 = 0{,}99975273$

$V_2 = 394{,}21 \times 10^{-6}$
$I_2 = 0{,}99960579$

TABLEAU VII

Evolution des coefficients des facteurs dans le cas du projet "Styles" (88 sujets)
(Méthode d'extraction: analyse en composantes principales, méthode de rotation: γ=0 (quartimax))

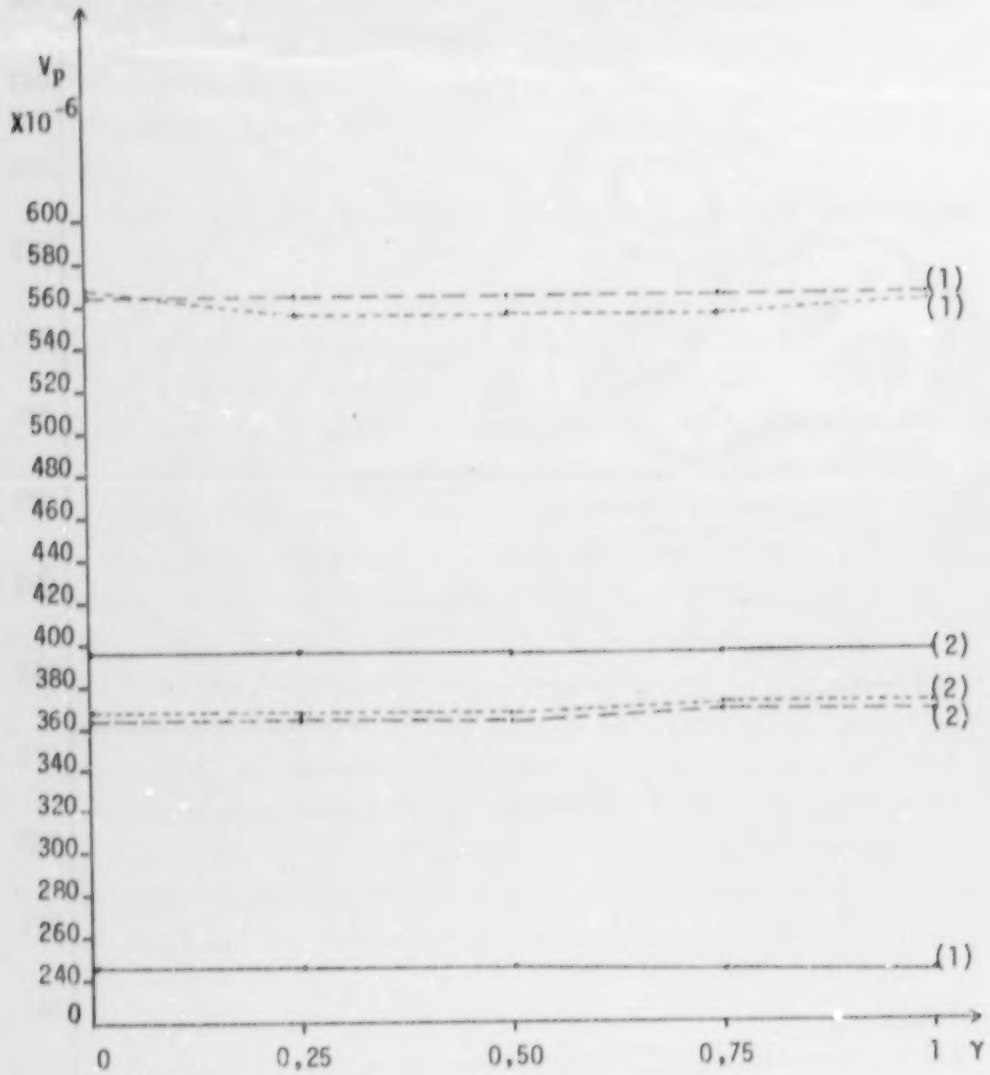
	Facteur 1					Facteur 2				
	4	6	8	10	12	4	6	8	10	12
V11	94	91	88	86	83	02	-05	-02	-03	-02
V21	02	05	02	03	02	92	90	86	87	86
V14	93	91	89	86	83	13	08	11	11	11
V23	12	18	14	15	14	90	83	78	77	75
V12		81	80	82	83		21	23	21	22
V24		00	-02	01	00		85	87	85	84
V16			82	79	80			-00	02	03
V25			11	11	09			72	69	71
V13				73	72				01	01
V22				-02	-02				45	41
V15					76					06
V26					34					53

$V_1 = 145{,}79 \times 10^{-6}$
$I_1 = 0{,}99985421$

$V_2 = 214{,}09 \times 10^{-6}$
$I_2 = 0{,}99978591$

TABLEAU VIII

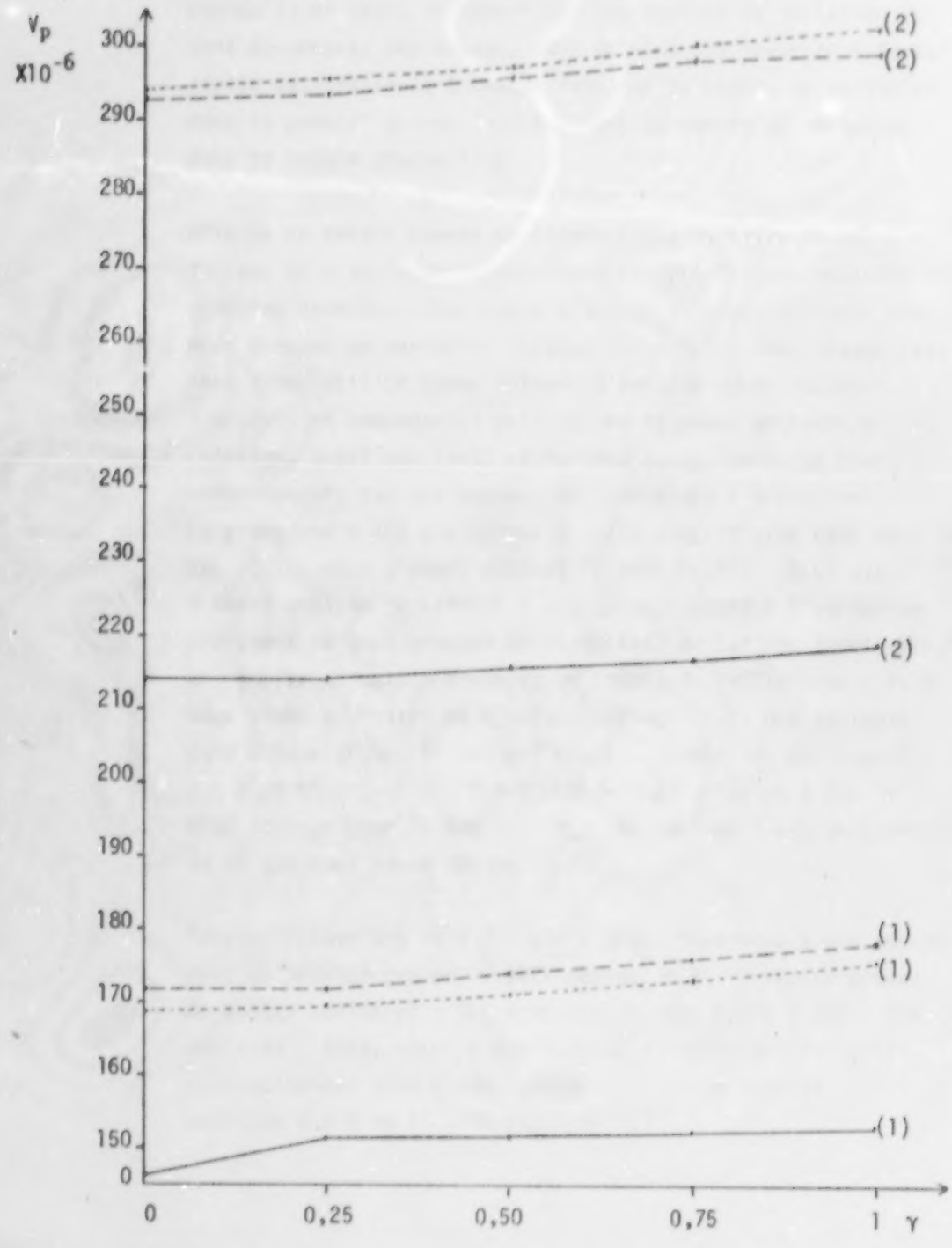
Courbes (approximées) de l'évolution de V_p en fonction de γ
(Projet "Styles d'apprentissage", cas 184 sujets)



Légende : ——— : PCA
– – – : PFA, Maxrow
----- : PFA, SMC
(1) : facteur 1
(2) : facteur 2

TABLEAU IX

Courbes (approximées) de l'évolution de V_p en fonction de γ
(Projet "Styles d'apprentissage", cas 88 sujets)



Légende : ——— : PCA
– – – : PFA, Maxrow
----- : PFA, SMC

(1) : facteur 1
(2) : facteur 2

Quoiqu'il en soit, il appert que les indices de variation Vp sont en général peu élevés, nous donnant une invariance factorielle satisfaisante à chaque fois que le nombre de variables dans le premier groupe (n_1) équivaut au nombre de variables dans le second groupe (n_2).

Afin de se rendre compte de l'importance relative de cette condition, $n_1 = n_2$, pour l'invariance factorielle des méthodes de rotation étudiées, nous avons effectué diverses analyses avec deux groupes de variables inégaux ($n_1 \neq n_2$). Dans chaque cas, nous avons utilisé comme méthode d'extraction de facteurs, l'analyse en composantes principales et comme méthodes de rotation, quartimax ($\gamma=0$) et varimax ($\gamma=1$). Nous ne nous sommes servis que des données du "Thurstone box problem". La procédure a été analogique à celle décrite plus haut dans le cas où les deux groupes avaient la même taille. Nous avons d'abord analysé la matrice 5 x 5 correspondant à 5 variables provenant de deux groupes de variables, de tailles distinctes, un premier de taille trois et un second de taille deux. Puis nous avons poursuivi en ajoutant à chaque fois une variable dans chaque groupe et analysé successivement les matrices 7 x 7, 9 x 9 et 11 x 11. L'ordre d'entrée des variables a été le même ici que pour le cas $n_1 = n_2$. Le tableau X est un exemple de ce que nous avons obtenu.

Pour effectuer une dernière série d'analyses nous avons débuté avec un premier groupe de deux variables et un second groupe de quatre variables. Une première matrice 6 x 6 a ainsi été analysée. Puis, comme d'habitude en ajoutant une variable successivement dans chaque groupe nous avons analysé les matrices 8 x 8 et 10 x 10 (tableau XI).

TABLEAU X

Evolution des coefficients des facteurs dans le cas de deux groupes de variables inégaux (δ=1)
(Méthode d'extraction: analyse en composantes principales, méthode de rotation: γ=1 (varimax))

	Facteur 1				Facteur 2			
	5	7	9	11	5	7	9	11
V11	08	05	01	14	97	98	94	98
V21	93	89	86	78	11	09	06	01
V13	30	30	27	45	93	97	97	88
V23	97	97	96	96	18	19	20	06
V26	92	95	93	98	26	35	37	19
V16		46	48	59		56	62	53
V22		94	94	99		27	31	13
V14			51	70			81	64
V24			88	97			38	17
V15				-53				56
V25				01				-22

$V_1 = 1437{,}71 \times 10^{-6}$
$I_1 = 0{,}99856229$

$V_2 = 2347{,}63 \times 10^{-6}$
$I_2 = 0{,}99765237$

Si on note $\delta = n_1 - n_2$, le tableau XII nous donne l'augmentation des coefficients de variation factorielle Vp lorsque δ passe de 0 à 1 à 2. On remarque immédiatement les différences notoires entre les coefficients suivant que l'analyse s'est effectuée avec des groupes de taille identique (δ=0), avec des groupes dont la différence de taille est 1 (δ=1), ou avec des groupes dont la différence de taille est 2 (δ=2). Prouvant par le fait même la pertinence de la condition $n_1 = n_2$ en rapport avec l'invariance factorielle.

CONCLUSION

Etant donné deux groupes G_1, G_2 de variables colinéaires nous avons démontré que la condition $n_1 = n_2$ suffisait pour que l'ensemble des méthodes de rotation orthomax normalisées donnent des facteurs invariants: ce résultat se situant dans le prolongement de celui obtenu par Kaiser (1958) pour le varimax normalisé. Par la suite, nous avons vérifié ce résultat empiriquement en nous servant de trois ensembles de deux groupes de variables. Comme prévu, l'analyse de ces trois ensembles de données a confirmé le résultat théorique dans le sens que les facteurs ainsi obtenus ont montré une invariance factorielle fort satisfaisante. Cependant on a pu se rendre compte, en effectuant des analyses sur des groupes de variables où $n_1 \neq n_2$, de l'accroissement très sensible du coefficient de variation factorielle Vp à mesure que δ ($=n_1-n_2$) augmentait. En effet, par exemple, le coefficient V2 passe de 86, 23 ($\times 10^{-6}$) lorsque δ=0, à 1316,39 pour δ=1 et enfin à 8158,38 pour δ=2, montrant à quel point l'invariance factorielle peut être modifiée en ajoutant une ou deux variables de plus dans un groupe. Ces résultats nous indiquent aussi l'importance d'un bon choix de variables lorsque l'on a l'intention d'entreprendre une analyse factorielle.

TABLEAU XI

Evolution des coefficients des facteurs dans le cas de deux groupes de variables inégaux (δ=2)
(Méthode d'extraction: analyse en composantes principales, méthode de rotation: γ=1 (varimax))

	Facteur 1			Facteur 2		
	6	8	10	6	8	10
V_{11}	08	03	35	98	93	92
V_{21}	90	87	77	10	06	04
V_{13}	32	30	63	93	95	71
V_{23}	98	97	95	17	19	-07
V_{26}	94	94	98	25	29	-05
V_{22}	95	95	99	23	28	-09
V_{14}		53	82		77	30
V_{24}		90	98		34	-10
V_{15}			-39			83
V_{25}			-03			14
	$V_1 = 5681,36\times10^{-6}$			$V_2 = 14020,34\times10^{-6}$		
	$I_1 = 0,99431864$			$I_2 = 0,98597966$		

TABLEAU XII

Augmentation des coefficients de variation factorielle V_p en fonction de δ

V_1

$\gamma = 0$

$\delta = 0$	$\delta = 1$	$\delta = 2$
57,64	cas 1: 668,20	cas 3: 2489,50
	cas 2: 221,88	cas 4: 237,91

$\gamma = 1$

$\delta = 0$	$\delta = 1$	$\delta = 2$
51,98	cas 1: 1437,71	cas 3: 5681,36
	cas 2: 129,15	cas 4: 208,79

V_2

$\gamma = 0$

$\delta = 0$	$\delta = 1$	$\delta = 2$
86,23	cas 1: 1316,39	cas 3: 8158,38
	cas 2: 365,26	cas 4: 932,18

$\gamma = 1$

$\delta = 0$	$\delta = 1$	$\delta = 2$
68,57	cas 1: 2347,63	cas 3:14020,34
	cas 2: 258,60	cas 4: 650,57

Notes: 1) Tous les nombres de ce tableau doivent être multipliés par 10^{-6}.

2) Le cas 1 correspond au cas où la variable V_{12} a été enlevée de l'analyse.
Le cas 2 correspond au cas où la variable V_{26} a été enlevée de l'analyse.
Le cas 3 correspond au cas où les variables V_{12} et V_{16} ont été enlevées de l'analyse.
Le cas 4 correspond au cas où les variables V_{22} et V_{26} ont été enlevées de l'analyse.

BIBLIOGRAPHIE

Ahmavaara, Y., *The mathematical theory of factorial invariance under selection.* Psychometrika 19, 27-38, 1954.

Bloxom, B., *Alternative approaches to factorial invariance.* Psychometrika 37, 425-440, 1972.

Carroll, J.B., *Biquartimin criterion for rotation to oblique simple structure in factor analysis,* Science 126, 114-15, 1957.

Harman, H.H., *Modern factor analysis.* The University of Chicago Press, 1976.

Hendrickson, A.E., White, P.O., *Promax: A quick method for rotation to oblique simple structure,* British Journal of mathematical and statistical psychology, 17, 65-70, 1964.

Hofmann, R.J. *The orthotran solution,* Multivariate behavioral research, 13, 90-124, 1978.

Horst, P., *Factor analysis of data matrices,* Holt, Rihenart & Winston, 1965.

Joreskog, K.G., *Simultaneous factor analysis in several populations.* Psychometrika 36, 409-426, 1971.

Kaiser, H.F., *The varimax criterion for analytical rotation in factor analysis,* Psychometrika, 23, 187-200, 1958.

Meredith, W., *Rotation to achieve factorial invariance.* Psychometrika 29, 187-206, 1964.

Mulaik, S.A., *The foundations of factor analysis.* New-York, McGraw-Hill, 1972.

Neuhaus, J.O., Wrigley, C., *The quartimax method: An analytical approach to orthogonal simple structure.* British journal of mathematical and statistical psychology, 7, 81-91, 1954.

Rummel, R.J., *Applied factor analysis.* Evanston: Northwestern University Press, 1970.

Saunders, D.R., *Trans-varimax,* American Psychologist, 17, 395, 1962.

Thurstone, L.L., *Multiple factor analysis.* Chicago, The University of Chicago Press, 1947.